

Söbke, Heinrich; Kämmerer, Frauke

Vermessene Fragen. Metriken als Ansatz automatisierter analytischer und konstruktiver Qualitätssicherung von Mehrfachauswahlfragen für mobile digitale Medien

Pfau, Wolfgang [Hrsg.]; Baetge, Caroline [Hrsg.]; Bedenlier, Svenja Mareike [Hrsg.]; Kramer, Carina [Hrsg.]; Stöter, Joachim [Hrsg.]: Teaching Trends 2016. Digitalisierung in der Hochschule: Mehr Vielfalt in der Lehre. Münster ; New York : Waxmann 2016, S. 153-162. - (Digitale Medien in der Hochschullehre; 5)



Quellenangabe/ Reference:

Söbke, Heinrich; Kämmerer, Frauke: Vermessene Fragen. Metriken als Ansatz automatisierter analytischer und konstruktiver Qualitätssicherung von Mehrfachauswahlfragen für mobile digitale Medien - In: Pfau, Wolfgang [Hrsg.]; Baetge, Caroline [Hrsg.]; Bedenlier, Svenja Mareike [Hrsg.]; Kramer, Carina [Hrsg.]; Stöter, Joachim [Hrsg.]: Teaching Trends 2016. Digitalisierung in der Hochschule: Mehr Vielfalt in der Lehre. Münster ; New York : Waxmann 2016, S. 153-162 - URN: urn:nbn:de:0111-pedocs-188957 - DOI: 10.25656/01:18895

<https://nbn-resolving.org/urn:nbn:de:0111-pedocs-188957>

<https://doi.org/10.25656/01:18895>

in Kooperation mit / in cooperation with:



WAXMANN
www.waxmann.com

<http://www.waxmann.com>

Nutzungsbedingungen

Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Die Nutzung stellt keine Übertragung des Eigentumsrechts an diesem Dokument dar und gilt vorbehaltlich der folgenden Einschränkungen: Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use

We grant a non-exclusive, non-transferable, individual and limited right to using this document.

This document is solely intended for your personal, non-commercial use. Use of this document does not include any transfer of property rights and it is conditional to the following limitations: All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

Kontakt / Contact:

peDOCS
DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation
Informationszentrum (IZ) Bildung
E-Mail: pedocs@dipf.de
Internet: www.pedocs.de

Mitglied der


Leibniz-Gemeinschaft

Wolfgang Pfau, Caroline Baetge,
Svenja Mareike Bedenlier, Carina Kramer,
Joachim Stöter (Hrsg.)

Teaching Trends 2016

Digitalisierung in der Hochschule:
Mehr Vielfalt in der Lehre



Waxmann 2016
Münster • New York

Bibliografische Informationen der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.dnb.de> abrufbar.

Digitale Medien in der Hochschullehre, Band 5

ISSN 2199-7667

ISBN 978-3-8309-3548-3

© Waxmann Verlag GmbH, Münster 2016
Steinfurter Straße 555, 48159 Münster

www.waxmann.com

info@waxmann.com

Umschlaggestaltung: Steffen Ottow, Clausthal-Zellerfeld

Titelbild: © kasto – fotolia.com

Satz: Sven Solterbeck, Münster

Druck: Hubert & Co., Göttingen

Gedruckt auf alterungsbeständigem Papier,
säurefrei gemäß ISO 9706



Printed in Germany

Alle Rechte vorbehalten. Nachdruck, auch auszugsweise, verboten.
Kein Teil dieses Werkes darf ohne schriftliche Genehmigung des
Verlages in irgendeiner Form reproduziert oder unter Verwendung
elektronischer Systeme verarbeitet, vervielfältigt oder verbreitet werden.

Inhalt

Grußwort der Niedersächsischen Ministerin für Wissenschaft und Kultur, Dr. Gabriele Heinen-Kljajić	9
---	---

Vorwort des Vorstandsvorsitzenden des ELAN e. V., Thomas Hanschke	11
--	----

Vorwort des Vizepräsidenten für Internationales, Weiterbildung und Digitalisierung der Technischen Universität Clausthal, Prof. Dr. Wolfgang Pfau	13
---	----

A Herausforderung Diversität: individualisiertes Lehren und Lernen mit digitalen Medien

<i>Martina Emke, Claudia Lehmann und Annelene Sudau</i> Die individuelle Online-Studienvorbereitung beruflich Qualifizierter auf dem OHN-KursPortal im Spannungsfeld zwischen Offenheit und Verbindlichkeit	17
--	----

<i>Anna Maria Schulz, Claudia Frie, Doris Meißner und Ralf Steffen</i> System capacity building Einsatz von Adobe Connect in der Entwicklungszusammenarbeit	27
---	----

<i>Nadine Schaarschmidt, Claudia Albrecht und Claudia Börner</i> Videoeinsatz in der Lehre Nutzung und Verbreitung in der Hochschule	39
--	----

<i>Linda Eckardt, Simone Kibler und Susanne Robra-Bissantz</i> Entwicklung eines Serious Games zum Lernen von Informationskompetenz und Leitlinien zur Nachnutzung	49
--	----

<i>André Schneider, Verena Jahn und Linda Heise</i> Digitalisierung der Lehre als Chance zur Ermöglichung dualer Karrieren studierender Spitzensportler	63
---	----

<i>Peter Ferdinand, Sergei Pachtchenko und Christian Schowalter</i> E-Studienvorbereitung mit personalisiert adaptierten E-Learning- Modulen und Tests zum leichteren Einstieg ins Studium an der Universität Koblenz-Landau	73
---	----

Jens Hilgedieck

Intercultural Learning Network

Erwerb interkultureller Kompetenzen durch virtuelle Mobilität 81

Nico Raichle und Kerstin Voß

„work&study“ – offene Hochschulen Rhein-Saar

Entwicklung polyvalenter Hochschulangebote für nichttraditionell

Studierende im Blended-Learning-Format 91

Christian-Henrik Walter

Gesamtkonzept zur Integration individualisierten Lernens

mittels Lehrvideos – ein Praxisbeispiel

Beispiel und Erfahrungsbericht zur Umsetzung eines Lehrkonzeptes

zum Einsatz integrierter Lehrvideos als zentrales Element einer

überarbeiteten Lehrveranstaltung 101

Mirjam Bretschneider und Ellen Pflaum

Lernendenzentrierung im Lehren und Lernen mit Medien 111

B Erfolgsfaktoren des Einsatzes digitaler Medien an Hochschulen

Katrin Schulenburg, Eva-Maria Beck, Wibke Hollweg,

Silke Trock, Elke Kraus und Theda Borde

Kollegiale Hospitation zur Stärkung des Einsatzes

neuer Medien in der Hochschullehre 123

Nicole Bellin-Mularski

Das (E-)Portfolio im Praxissemester:

Dokumentations- oder Reflexionsinstrument? 131

Birte Heidkamp und David Kergel

Rückblick und Ausblick

Das mobile E-Learning-Center der Carl von Ossietzky Universität

Oldenburg und das E-Learning-Zentrum der Hochschule Rhein-Waal

im Kontext der Digitalisierung der Lehre 143

Heinrich Söbke und Frauke Kämmerer

Vermessene Fragen

Metriken als Ansatz automatisierter analytischer und konstruktiver

Qualitätssicherung von Mehrfachauswahlfragen für mobile digitale Medien .. 153

Heinrich Söbke und Maria Reichelt

„Rat(t)en in der Lehre“

Über die Spiel(un)lust unserer Studierenden am Beispiel digitaler Apps 163

Lars Rettig und Tim Warszta

Der Einfluss von Kursdesignelementen auf Studierendenzufriedenheit
und Studierendenloyalität

Ein Policy-Capturing-Design-Ansatz 177

Kerstin Voß und Nico Raichle

Anreize, Motivation und Support für Lehrende zum
Einsatz digitaler Medien in der Hochschullehre

Ergebnisse einer qualitativen Expertenbefragung aus dem Verbundprojekt
„work&study“ – offene Hochschulen Rhein-Saar 191

Sulamith Frerich und Silke Frye

Erfolgreiches Praxisbeispiel

Ein interaktives Blended-Learning-Seminar mit Praxisphase 199

Jana Riedel und Claudia Börner

Wir tun es, weil es gut ist!

Wie Lehrende die Erfolgsfaktoren für den Einsatz digitaler Medien
in der Hochschullehre einschätzen 209

Nico Raichle

Quests, Raids. Level Up ... Game Over?!

Erfolgsfaktoren von Gamification in der Hochschullehre 221

Autorinnen und Autoren 235

Vermessene Fragen

Metriken als Ansatz automatisierter analytischer und konstruktiver Qualitätssicherung von Mehrfachauswahlfragen für mobile digitale Medien

Abstract

Auswahlfragen (Multiple-Choice-Fragen) werden in vielen Lehr-Lern-Kontexten eingesetzt. Insbesondere in Verbindung mit digitalen Medien sind Auswahlfragen von hoher Bedeutung, da sie zum einen Lernende durch notwendige Entscheidungen mit nachfolgender Interaktion aktivieren und zum anderen die Möglichkeit einer automatisierbaren Erfolgskontrolle gegeben ist. Insbesondere durch die in den letzten Jahren aufgekommenen Möglichkeiten des mobilen Lernens mit Hilfe von Apps haben sich neue Einsatzkontexte für Auswahlfragen ergeben. Die vermehrten Einsatzmöglichkeiten machen die weitere Erstellung von Auswahlfragen notwendig. Die Komplexität des Entwurfs guter Auswahlfragen wird jedoch häufig unterschätzt. Deren Qualität ist aber ein erfolgskritischer Faktor des Einsatzes von digitalen Medien. Daher haben wir bestehende Entwurfsrichtlinien gesammelt. Zusätzlich entstehen durch neue Medien (Apps) zusätzliche Anforderungen. In eigenen empirischen Untersuchungen in ingenieurwissenschaftlichen Studiengängen konnten wir feststellen, dass kurze, prägnante Fragen und Antworten aufgrund geringerer Bildschirmgrößen und App-induziertem Zeitdruck von hoher Bedeutung für die Benutzbarkeit sind. Wir schlagen daher die Nutzung von Metriken zur objektiven Messung der textuellen Komplexität von Auswahlfragen vor. Derartige Metriken sind automatisiert anwendbar und daher ein Schritt in Richtung einer maschinellen Qualitätssicherung für Auswahlfragen. Zusätzlich lassen sie sich zur Ableitung von Entwurfsrichtlinien im Sinne einer konstruktiven Qualitätssicherung nutzen.

Einleitung

Auswahlfragen (oder Multiple Choice Questions – MCQ) sind ein wichtiges Werkzeug in der Lehre. Zu ihrer großen Bedeutung tragen mehrere Faktoren bei. Sie sind nicht auf bestimmte Wissensgebiete und -arten beschränkt, sondern vielmehr universell einsetzbar. Zusätzlich werden Lernprozesse – im einfachsten Fall die Speicherung und der Abruf von Faktenwissen, jedoch in komplexeren Fällen auch die Anwendung von Wissen höherer Ebenen, z.B. mit der Präsentation eines für die Lernenden unbekannten Problems in der Auswahlfrage (Haladyna & Rodriguez, 2013; Iz & Fok, 2007) – durch die implizit notwendige Interaktivität (die Auswahl der korrekten Lösung) unterstützt. Darüber hinaus ist eine automatisierte Überprüfung – und damit aufwandsarme Korrektur – der korrekten Lösung möglich. Dies ist insbesondere aufgrund des Einsatzes in digitalen Medien von steigender

Bedeutung. Außerdem ergeben sich größere Einsatzspektren und -ziele. Zum einen werden MCQs seit langer Zeit im Rahmen der Lernzielkontrolle eingesetzt. Zusätzlich wurde auch nachgewiesen, dass MCQs und ihre wiederholte Ausführung als Lernwerkzeug geeignet sind (Agarwal, Karpicke, Kang, Roediger & McDermott, 2008; Karpicke & Roediger, 2008; Pashler et al., 2007). Im Rahmen dieser Arbeit gehen wir davon aus, dass Lernen mehrere komplexe Prozesse umfasst und wohl das wichtigste Ziel formaler Bildungskontexte darstellt. Ein Prozess ist der Erwerb von deklarativem Wissen („Wissen, dass“). Deklaratives Wissen kann einzelne Fakten umfassen (z. B. eine mathematische Formel), aber auch Wissen über komplexe Zusammenhänge (z. B. Verständnis von Wechselwirkungen im Wasserkreislauf) beinhalten, bei dem ein tiefes Verständnis von Konzepten und deren Wechselwirkungen aufgebaut werden muss (Renkl, 2015). Erwerb und Abruf deklarativen Wissens kann mit MCQs unterstützt werden. Gleichfalls wird im Rahmen eines konstruktivistischen Lernansatzes (Papert, 1991) die Erstellung von MCQs als unterstützend für eine tiefergehende Beschäftigung mit den Lerninhalten angesehen (Parker, Manuguerra & Schaefer, 2013). Außerhalb der Lehre erfreuen sich MCQs in Form von Quizzes großer Beliebtheit und scheinen ein großes Faszinationspotenzial zu besitzen, wie beispielsweise die Popularität sowohl von Spielen (z. B. Gesellschaftsspiel Trivial Pursuit (Bellis, o. J.)) und Spiel-Apps (z. B. Quizduell (FEO Media AB, 2013) und QuizUp (Plain Vanilla, 2014)) sowie Fernsehsendungen (z. B. Wer wird Millionär? (RTL Television, 1999)) zeigt.

Für all diese Anwendungsfälle ist die Erstellung geeigneter Fragen keineswegs als trivial anzusehen. Sie ist daher Gegenstand wissenschaftlicher Untersuchungen (Burton, 2005; Campbell, 2011; Haladyna, Downing & Rodriguez, 2002; Haladyna & Rodriguez, 2013; Iz & Fok, 2007; McAlpine & Hesketh, 2003; McCoubrie, 2004). Insbesondere die Erstellung guter Distraktoren – die nicht zutreffenden Antwortoptionen – ist nicht trivial. Da Fragen von den Lehrenden und Lernenden häufig selbst entwickelt werden müssen, sind Leitfäden für die zielorientierte Erstellung von MCQs notwendig. Diese zeigen nach unseren Erfahrungen positive Wirksamkeit (Söbke & Parker, 2016; Söbke & Weitze, 2016). Mit dem Aufkommen neuer digitaler Medien können sich die Anforderungen an Fragen ändern. Beispiel hierfür sind Apps, die auf mobilen Endgeräten für eine eher beiläufige Beschäftigung mit den Fragen sorgen (Micro-Learning, Lernen in Leerzeiten, ÖPNV (Reichelt & Söbke, 2016)). Zu den relevanten Unterschieden gegenüber den bisher für den Einsatz von MCQs genutzten Medien zählen der begrenzte Platz auf den Displays mobiler Geräte sowie die tendenziell kürzeren Lernzeiten beim mobilen Lernen.

In diesem Artikel geben wir einen Überblick über essentielle Gestaltungsgrundsätze von Auswahlfragen. Der Gestaltungsgrundsatz von kurzen und prägnanten Fragen ist insbesondere für mobile Medien von Bedeutung. Daher gehen wir am Beispiel dieses Grundsatzes der Forschungsfrage nach, inwieweit sich die notwendige Qualitätssicherung automatisieren lässt. Zur Umsetzung schlagen wir die Nutzung von Metriken im Rahmen einer analytischen Qualitätssicherung vor und zeigen in

einem Beispiel deren prinzipielle Validität sowie die daraus mögliche Ableitung von Entwurfsrichtlinien als Maßnahme der konstruktiven Qualitätssicherung.

Grundsätze zur Gestaltung von Auswahlfragen

Aufgrund der enormen Bedeutung von Auswahlfragen für die Lehre sind bereits zahlreiche Handreichungen für deren Entwicklung verfügbar. In diesem Abschnitt fassen wir die existierenden Grundsätze zur Erstellung guter Auswahlfragen zusammen. Insbesondere die folgenden Richtlinien wurden immer wieder genannt (z. B. Brauns & Schubert, 2008; Feger, 1984; Haladyna & Rodriguez, 2013; Krüger, 2013):

- Es sollte relevanter Inhalt abgefragt werden (d. h. nicht zu allgemein, aber auch nicht zu spezifisch).
- Die linguistische Komplexität sollte an die Zielgruppe angepasst sein.
- Die Länge von Frage und Antworten (Leseaufwand) sollte kurzgehalten werden.
- Die grundlegende Hauptinformation soll in der Frage gegeben werden.
- Die Frage sollte positiv formuliert sein.
- Die Frage sollte so formuliert sein, dass sie auch ohne Kenntnis der Antwortoptionen beantwortet werden kann.
- Es kann mit Abstufungen gearbeitet werden (z. B. einfachste, schnellste).
- Es sollten keine absichtlichen Fallen in den Fragen sein.

Besonders die Erzeugung von plausiblen Distraktoren ist eine anspruchsvolle Aufgabe, da unwahrscheinliche Antwortoptionen die Auswahl der richtigen Antwort erleichtern. Die folgenden Grundsätze sollen die Findung glaubhafter Distraktoren unterstützen:

- Die Antworten sollten voneinander unabhängig sein.
- Zahlenbereiche sollten sich nicht überlappen.
- Länge, Stil und Grammatik sollten sich an der richtigen Antwort orientieren.
- Antworten müssen zweifellos falsch sein.
- Antworten sollten plausibel und von typischen Fehlern abgeleitet sein.
- Abkürzungen sollten vermieden werden.
- Antwortoptionen sollten alphabetisch oder numerisch sortiert angeordnet werden.
- Es sollten keine Hinweise auf die richtige Antwort gegeben werden. Dazu zählen Worte wie nie, immer oder vollkommen in den Distraktoren, die Wiederholung eines Wortes aus der Frage in der richtigen Antwort und offensichtlich falsche Distraktoren.

Bemerkenswerterweise argumentieren Haladyna & Rodriguez (2013), dass insgesamt drei Antwortoptionen für eine Bewertung des Wissenstandes ausreichend

sind. Diese Aussage kollidiert mit den programmatischen Möglichkeiten einiger Testwerkzeuge, senkt jedoch den Aufwand beim Fragenentwurf.

Sowohl bei diesen Richtlinien als auch den vorherigen Richtlinien zur Findung geeigneter Fragen ist auffällig, dass sie eher weichen Kriterien entsprechen. Die meisten von ihnen sind nicht objektiv validierbar. Mit dem folgenden Vorschlag möchten wir dieses zumindest teilweise ändern.

Andere Anforderungen durch neue Medien: mobile Fragen-Apps

Der Popularitätsschub von kommerziellen Quiz-Apps 2013/2014 mit den Marktführern Quizduell und QuizUp, die jeweils mehrere Millionen Installationen in Deutschland erreichen konnten, hat sich das Format Quiz auf mobilen Endgeräten etabliert. Abgesehen von der Möglichkeit, QuizUp für selbstdefinierte Fragensätze zu nutzen (Woods, 2015), gibt es gleichfalls eine Reihe von speziellen, zu Lernzwecken entwickelte Apps. Dazu gehören beispielsweise Quizlet (Quizlet LLC, 2015), StudyBlue (StudyBlue Inc., 2016), Skive (qLearning Applications GmbH, 2016) und BrainYoo (BrainYoo Ltd., 2015). Ebenfalls arbeiten führende Lernplattformen mit sogenannten *Responsive Designs*, d.h. die enthaltenen MCQ-Module eignen sich auch für die Bearbeitung auf mobilen Endgeräten. Insgesamt gesehen stellen diese Technologien ein enormes Potenzial für die Beantwortung von MCQs auf mobilen Endgeräten dar. Gegenüber bisher genutzten Medien ergibt sich insbesondere die Anforderung, dass das Platzangebot auf den Displays mobiler Endgeräte begrenzt ist. Daraus folgt die Notwendigkeit kurzer und prägnanter Fragen und Antworten. Abbildung 1 zeigt ein Beispiel zweier nicht optimaler MCQs und ihrer Auswirkungen auf das Erscheinungsbild. Weiterer Nachdruck wird dieser Anforderung durch eine mögliche Zeitbeschränkung für die Wahl der Antwort verliehen, wie er in Quiz-Apps üblich ist. Je kürzer der zu lesende Text ist, desto mehr Zeit verbleibt für den Antwortprozess. Eine Nichtbeachtung dieser Anforderung reduziert die Lerneffizienz, da teilweise aufgrund des Zeitdrucks nach optischen Merkmalen ausgewählt wird. Das heißt, hier findet dann eine optische Zuordnung von Antwort zu Frage statt – es findet keine semantische Auseinandersetzung mit der MCQ statt. Zudem ergaben die Untersuchungen die noch zu bestätigende Hypothese, dass eine hohe textuelle Komplexität der Frage gleichfalls der Akzeptanz abträglich ist (Söbke & Weitze, 2016). Gefolgert werden kann die Notwendigkeit von Fragen, die sich (kognitiv) möglichst schnell erfassen lassen und gleichzeitig dennoch die mit der Frage beabsichtigten Lernziele (z. B. Speicherung und Abruf einer mathematischen Formel oder Anwendung dieser Formel auf ein mathematisches Problem) unterstützen.

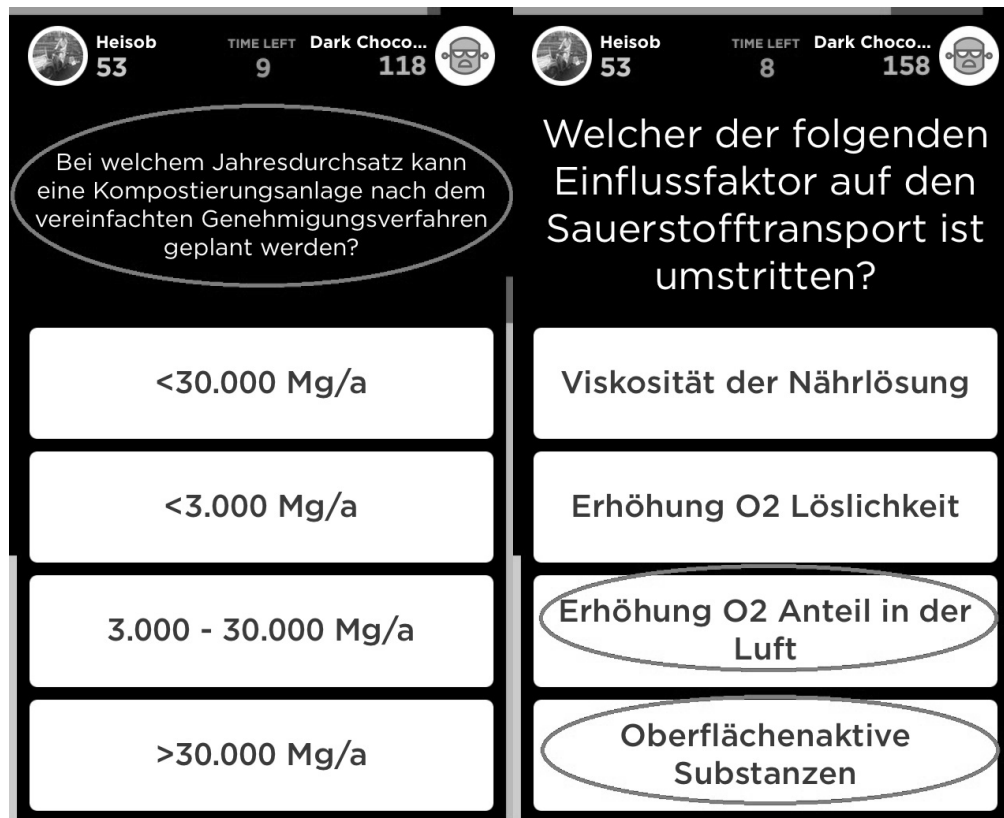


Abb. 1: Auswirkungen langer Zeichenfolgen für Frage (links) und Antwort (rechts) am Beispiel der App QuizUp und Fragen aus dem Bildungsthema „Luftikusse“

Merkmale kognitiv einfach zu erfassender Fragen

Bei der Unterteilung der Maßnahmen zum Entwurf von einfach zu erfassenden Fragen kann eine Makro- und eine Mikroebene unterschieden werden. Zur Makroebene zählen wir alle Maßnahmen, die über eine einzelne Frage hinaus die Verständlichkeit einer Gruppe von Fragen verbessert. Dazu zählt beispielsweise ein entsprechendes Level-Design, d. h. eine gruppenweise Unterteilung der Fragen, die es erlaubt, Gruppen von Fragen nach und nach zur Bearbeitung freizugeben, um aufeinander aufbauendes Wissen zu vermitteln. Eine solche Unterteilung kann die Motivation erhöhen, da die Lernenden dann keine Fragen beantworten müssen, die sie noch nicht in einen Gesamtkontext einordnen können (Söbke & Weitze, 2016).

Auf der Mikroebene hingegen werden die Maßnahmen zum Entwurf einer einzelnen Frage eingeordnet. In dieser Ebene beschäftigen wir uns im Folgenden mit der textuellen Komplexität einer Frage. Der obigen Hypothese zufolge sollte diese möglichst niedrig sein, um eine möglichst geringe wahrgenommene Anstrengung beim Lesen bzw. Erfassen einer Frage zu erreichen. Für die noch vorzunehmende Validierung ist eine Metrik notwendig, mit der die textuelle Komplexität einer Frage bestimmt werden kann.

Ähnliche Herausforderungen sind im Bereich der Softwareentwicklung zu lösen: Für die Wartungsfreundlichkeit eines Programmquelltextes ist eine niedrige

textuelle Komplexität von Vorteil. Dann ist der Quelltext für einen Softwareentwickler leichter verständlich und gezielte Änderungen werden einfacher. Vor diesem Hintergrund wurden verschiedene Metriken entwickelt, die der Messung der kognitiven Komplexität eines Quellcodetextes dienen. Beispielhaft genannt seien hier die sogenannten Halstead-Metriken (Halstead, 1977). Hier werden mehrere Basismaße benutzt: die Anzahl der unterschiedlichen Operatoren (η_1), die Anzahl der unterschiedlichen Operanden (η_2), die gesamte Anzahl der Operatoren (N_1) sowie die gesamte Anzahl der Operanden (N_2). Aus diesen Basismaßen lassen sich dann verschiedene weitere Kennzahlen berechnen. So zeigt beispielsweise Formel 1 die Berechnung der *Schwierigkeit* – eine Maßzahl für den Aufwand, einen Programmquelltext zu schreiben bzw. zu verstehen.

Formel 1: Schwierigkeit D nach Halstead

$$D = \frac{\eta_1}{2} \times \frac{N_2}{\eta_2}$$

Für eine Übertragung dieser Metrik auf einen Fragetext ist es notwendig festzulegen, wie die Begriffe Operator und Operand in Bezug auf den Fragentext definiert sind. Wir haben die in Tabelle 1 gezeigte Festlegung benutzt.

Tab. 1: Zuordnung grammatikalischer Elemente zu Halstaed-Basisgrößen

Basisgröße	Grammatikalische Elemente
Operator	Adjektiv, Adverb, Konjunktion, Präposition, Satzzeichen, Verb
Operand	Artikel, Nomen, Pronomen

Mit Hilfe dieser Festlegung ist es nun möglich, die Metrik auf eine Beispielfrage anzuwenden. Wir haben dazu die Beispielfrage aus Abbildung 1 (links) herangezogen und einen Alternativfragentext erstellt (vgl. Tabelle 2, Variante 2). Beim Entwurf dieses Textes haben wir auf eine kompakte sprachliche Darstellung Wert gelegt. Das führte dazu, dass der Ergebnistext jetzt keine grammatikalisch korrekte Frage mehr darstellt. Jedoch hat er im Kontext der Auswahlfrage nicht an Aussagekraft verloren. Die weiteren Spalten drei bis sechs der Tabelle 2 zeigen die Basismaße, wie wir sie gemäß der Zuordnung in Tabelle 1 bestimmt haben. Es ist erkennbar, dass Variante 2 deutlich niedrigere Basismaße hat als Variante 1 und dass dadurch auch die Schwierigkeitskennzahl D deutlich abgenommen hat. Damit konnte mehrere Zusammenhänge demonstriert werden. Erstens ist es prinzipiell mit Hilfe einer geeigneten Metrik möglich, rechnerisch – und damit potentiell automatisiert – eine Aussage über die textuelle Komplexität einer Frage zu treffen. Darüber hinaus ist anzumerken, dass aus der Metrik Konstruktionsprinzipien für die Fragen abgeleitet werden können. Eine Analyse z. B. der Formel 1 zeigt, dass sich der Schwierigkeitswert D verringern lässt, wenn entweder die Anzahl der unterschiedlichen Operato-

ren oder die Gesamtanzahl der Operanden verringert wird oder aber die Anzahl der unterschiedlichen Operanden – begrenzt durch deren Gesamtzahl – erhöht wird.

Tab. 2: Varianten des Fragentextes: Basismaße und Schwierigkeitsmetrik

Variante	Fragentext	η_1	η_2	N_1	N_2	D
1	„Bei welchem Jahresdurchsatz kann eine Kompostierungsanlage nach dem vereinfachten Genehmigungsverfahren geplant werden?“	7	6	7	6	3,5
2	„Jahresdurchsatz einer Kompostierungsanlage nach vereinfachtem Genehmigungsverfahren?“	3	4	3	4	1,5

Das vorgeschlagene Verfahren, die textuelle Komplexität einer Frage mit Hilfe von Metriken zu messen, kann bisher nur als aussichtsreicher Ansatz betrachtet werden. Die entsprechenden Basismaße müssen sicherlich noch erweitert werden. Kandidaten hierfür sind z.B. die Anzahl der Fremdwörter (erhöhen Komplexität), Anzahl der Gesamtzeichen (je mehr Zeichen, desto mehr muss gelesen werden) und separate Ausweisung von Satzzeichen als Basismaß (Satzzeichen sind als strukturalisierend zu werten). Gleichfalls müssen die Antwortoptionen in die Metrik einbezogen werden. Hier sind dann weitere Basismaße bzw. Metriken zu erheben, die beispielsweise die unterschiedliche Länge und Struktur der Antwortoptionen bewerten. An dieser Stelle ist es neben der Verringerung der Gesamtkomplexität ein Ziel, für alle Antwortoptionen möglichst gleichwertige textuelle Komplexität zu erreichen. Damit kann der Heuristik entgegengewirkt werden, dass die komplexeste Antwortoption meist die korrekte Lösung ist.

Mit dem vorgeschlagenen Ansatz ist es möglich, zum einen die textuelle Komplexität einer Frage automatisiert zu bewerten. Damit kann jeder Frage eine Maßzahl zugeordnet werden, die zumindest ein formales Qualitätsmaß darstellt. Zum anderen lassen sich aus den verwendeten Metriken auch Konstruktionsprinzipien ableiten. Kritisch anzumerken ist, dass beim demonstrierten Beispiel noch die rechnergestützte Kategorisierung der einzelnen Worte einer Frage fehlt. Zur Lösung dieser Aufgabe sind die (computer-)linguistischen Methoden zu nutzen. Zu prüfen ist gleichfalls, inwieweit diese Disziplin zu Basismaßen und Metriken beitragen kann. Festzustellen ist zudem, dass die textuelle Komplexität nur ein Qualitätskriterium einer MCQ ist. Die weiteren Kriterien, wie Verständlichkeit und Abdeckung der Lernziele, sind ebenfalls zu beachten. Diese bleiben weiter von Bedeutung, an ihrer automatisierten Messbarkeit muss gleichfalls gearbeitet werden.

Schlussfolgerungen und Ausblick

Die Bedeutung von Auswahlfragen als etabliertes Werkzeug für Lehr-Lern-Kontexte ist durch die Verbreitung von mobilen Endgeräten weiter gestiegen. Die damit verbundenen digitalen Lernformate bevorzugen kompakte Auswahlfragen. Die Richtlinien für den Entwurf von Auswahlfragen sind daher als erfolgskritisch für den Einsatz von digitalen Medien auch in der Hochschullehre zu sehen. In diesem Beitrag haben wir basierend auf einer Zusammenfassung von Entwurfsrichtlinien diese ergänzt um einen automatisierbaren Ansatz zur Messung der textuellen Komplexität im Rahmen der analytischen Qualitätssicherung. Diese Erweiterung bedarf jedoch noch weiterer Untersuchungen. Zum einen ist ein konkretes Modell zu entwerfen und dessen Wirksamkeit in Bezug auf die Lerneffizienz zu überprüfen. Des Weiteren wurden bisher mögliche zusätzliche mediale Elemente (wie Grafiken und Audioelemente) einer Auswahlfrage nicht betrachtet. Neben einer Steigerung der Prägnanz der Fragen – d. h. mit einer geringen textuellen Komplexität exakt und zielgerichtet zu fragen – sind weitere visionäre Ziele des Einsatzes von Auswahlfragen eine durchgängig automatisierte Qualitätssicherung von Fragen, ggf. mit den Methoden des *machine learning* (z. B. Anderka, Stein & Lipka, 2011) sowie die automatisierte, curriculumbezogene Generierung von Fragen, die zu einer vollständigen Abdeckung der Lernziele führt und deren manueller Ansatz beispielsweise durch Feger (1984) beschrieben wird.

Literatur

- Agarwal, P. K., Karpicke, J. D., Kang, S. H. K., Roediger, H. L. & McDermott, K. B. (2008). Examining the testing effect with open- and closed-book tests. *Applied Cognitive Psychology*, 22 (7), 861–876. John Wiley & Sons. Ltd. doi:10.1002/acp.1391.
- Anderka, M., Stein, B. & Lipka, N. (2011). Detection of text quality flaws as a one-class classification problem (CIKM '11). In *Proceedings of the 20th ACM international conference on Information and knowledge management* (S. 2313–2316). New York, NY, USA: ACM. doi:10.1145/2063576.2063954.
- Bellis, M. (o. J.). *The History of Trivial Pursuit*. Verfügbar unter: http://inventors.about.com/library/inventors/bl_trivia_pursuit.htm [13.1.2015].
- BrainYoo Ltd. (2015). *Brainyoo Karteikarten-App & Lernsoftware*. Verfügbar unter: <https://www.brainyoo.de/> [23.6.2016].
- Brauns, K. & Schubert, S. (2008). Qualitätssicherung von Multiple-Choice-Prüfungen. In S. Dany, B. Szcyrba & J. Wildt (Hrsg.), *Prüfungen auf die Agenda – Hochschuldidaktische Perspektiven auf Reformen im Prüfungswesen* (S. 92–102). Bielefeld: W. Bertelsmann Verlag GmbH & Co KG.
- Burton, R. F. (2005). Multiple-choice and true/false tests: myths and misapprehensions. *Assessment & Evaluation in Higher Education*, 30 (1), 65–72. doi:10.1080/0260293042003243904.
- Campbell, D.E. (2011). How to write good multiple-choice questions. *Journal of paediatrics and child health*, 47 (6), 322–5. doi:10.1111/j.1440-1754.2011.02115.x.

- Feger, B. (1984). Die Generierung von Testitems zu Lehrtexten. *Diagnostica*, 30 (1), 24–46.
- FEO Media AB. (2013). *QuizClash | Challenge your friends!* Zugriff am. Verfügbar unter: <http://www.quizclash-game.com/> [15.1.2015].
- Haladyna, T. M., Downing, S. M. & Rodriguez, M. C. (2002). A Review of Multiple-Choice Item-Writing Guidelines for Classroom Assessment. *Applied Measurement in Education*, 15 (3), 309–333. doi:10.1207/S15324818AME1503_5.
- Haladyna, T. M. & Rodriguez, M. C. (2013). *Developing and Validating Test Items*. New York: Routledge.
- Halstead, M.H. (1977). *Elements of Software Science* (Operating and Programming Systems Series). New York, NY, USA: Elsevier Science Inc.
- Iz, H.B. & Fok, H.S. (2007). Use of Bloom's taxonomic complexity in online multiple choice tests in Geomatics education. *Survey Review*, 39 (305), 226–237. doi:10.1179/003962607X165195.
- Karpicke, J. D. & Roediger, H. L. (2008). The critical importance of retrieval for learning. *Science* (New York, N.Y.), 319 (5865), 966–8. doi:10.1126/science.1152408.
- Krüger, M. (2013). *Erstellen und Bewerten von Multiple-Choice-Aufgaben*. Universität Hannover. Verfügbar unter: http://www.uni-hannover.de/imperia/md/content/elearning/practicalguides2/didaktik/elsa_handreichung_zum_erstellen_und_bewerten_von_mc_fragen.pdf.
- McAlpine, M. & Hesketh, I. (2003). Multiple response questions—allowing for chance in authentic assessments. In J. Christie (Hrsg.), *7th International CAA Conference*. Loughborough: Loughborough University.
- McCoubrie, P. (2004). Improving the fairness of multiple-choice questions: a literature review. *Medical teacher*, 26 (8), 709–12. doi:10.1080/01421590400013495.
- Papert, S. (1991). Situating Constructionism. In S. Papert & I. Harel (Hrsg.), *Constructionism*. Norwood, N.J.: Ablex Publishing.
- Parker, R., Manuguerra, M. & Schaefer, B. (2013). The Reading Game – encouraging learners to become question-makers rather than question-takers by getting feedback, making friends and having fun. In H. Carter, M. Gosper & J. Hedberg (Hrsg.), *30th ascilite Conference 2013 Proceedings* (S. 681–684). Macquarie University.
- Pashler, H., Bain, P.M., Bottge, B. A., Graesser, A., Koedinger, K., McDaniel, M. et al. (2007). *Organizing Instruction and Study to Improve Student Learning*. IES Practice Guide. NCER 2007–2004. National Center for Education Research. ERIC.
- Plain Vanilla. (2014). *QuizUp – Connecting people through shared interests*. Plain Vanilla. Verfügbar unter: <https://www.quizup.com/> [12.1.2016].
- qLearning Applications GmbH (2016). *Skive*. Verfügbar unter: <https://goskive.com/> [13.03.2016].
- Quizlet LLC. (2015). *Simple free learning tools for students and teachers | Quizlet*. Verfügbar unter: <http://quizlet.com/> [16.2.2015].
- Reichelt, M. & Söbke, H. (2016). „Spielend Lernen“ – Ein multimethodischer Blick auf die didaktische Anwendung einer Mobile Learning App im Bildungsraum Hochschule. In S. Krumme (Hrsg.), *Räume für Bildung. Räume der Bildung – DGfE-Kongress 2016*. Universität Kassel: Institut für Erziehungswissenschaft und Institut für Sozialwesen.
- Renkl, A. (2015). Wissenserwerb. In E. Wild & J. Möller (Hrsg.), *Pädagogische Psychologie* (S. 3–24). Berlin, Heidelberg: Springer Berlin Heidelberg. doi:10.1007/978-3-642-41291-2_1.
- RTL Television (1999). *Wer wird Millionär?* Verfügbar unter: <http://www.rtl.de/cms/sendungen/wer-wird-millionaer.html> [23.5.2016].

- Söbke, H. & Parker, R. (2016). *A Quiz Pipeline: Multifunctional Combination of Multiple Choice Question Based Games in a Formal Learning Context* (submitted).
- Söbke, H. & Weitze, L. (2016). The Challenge to Nurture Challenge – Students’ Perception of a Commercial Quiz App as a Learning Tool. In G. Wallner, S. Kriglstein, H. Hlavacs, R. Malaka & A. Lugmayr (Hrsg.), *Entertainment Computing – ICEC 2016–15th International Conference, Vienna, Austria, September 28–30, 2016, Proceedings* (Band 9926, S. 15–23). Springer International Publishing. doi:10.1007/978-3-319-46100-7_2.
- StudyBlue Inc. (2016). *Studyblue – Conquer Your Course*. Verfügbar unter: <https://www.studyblue.com/>.
- Woods, B. (2015, September 24). *QuizUp launches tools for creating your own trivia categories and questions*. *The Next Web*. Verfügbar unter: <http://thenextweb.com/apps/2015/09/24/quizup-launches-tools-for-creating-your-own-trivia-categories-and-questions/#gref> [12.1.2016].